

# Медицинская и биологическая кибернетика

УДК 004.9:314.303

О.А. Кривова, А.С. Коваленко, И.А. Курило

## СТАТИСТИЧЕСКАЯ ОЦЕНКА ВЛИЯНИЯ ФАКТОРОВ РЕГИОНАЛЬНОГО РАЗВИТИЯ НА РОЖДАЕМОСТЬ НАСЕЛЕНИЯ УКРАИНЫ

Разработана методика выделения демографических региональных структур на примере анализа рождаемости населения Украины и формирующих ее факторов в 1999–2006 гг., в которой используется кластерный, факторный анализ, а также индексный метод обоснования качества кластеризации. Проанализирована ситуация с рождаемостью в типологических регионах и ее изменение под влиянием выделенных комплексных факторов (материальное положение, урбанизация, неравенство, женская занятость). Дана статистическая оценка влияния основных факторов на уровень рождаемости.

### Введение

Взаимосвязь демографических и социально-экономических процессов давно является предметом теоретических построений и эмпирических исследований. В классических работах Г. Беккера, Т. Шульца и Р. Истерлина утверждается, что основным фактором, формирующим репродуктивное поведение, является доход семьи, в котором концентрируется широкий спектр социально-экономических условий воспроизводства населения [1–4]. Соотношением экономической и социальной детерминант авторы объясняют модернизацию современной рождаемости в развитых странах [5, 6]. Появились доказательства неоднозначности влияния материального благополучия на рождаемость [7]. Как известно, резкое снижение рождаемости в постсоветских странах явилось реакцией населения на падение уровня жизни, вызванное социально-экономическими трансформациями середины 90-х гг. [8–10]. Демографами изучены закономерности изменения рождаемости в годы кризиса и выхода из него [11–22]. Проводятся исследования демографических процессов на региональном уровне [17–19]. Проанализирована региональная дифференциация демографических процессов на территории России [17, 22], Украины [19, 20]. Выявлены социально-экономические факторы разного уровня, детерминирующие репродуктивное поведение населения в условиях трансформации общественных отношений [14, 15, 21–23].

В современных условиях социально-экономических рисков, экономической нестабильности актуальным является комплексный анализ развития регионов Украины, в котором бы учитывалось взаимовлияние демографических и социально-экономических процессов. Разработка статистических моделей дифференциации демографических процессов в

зависимости от социально-экономических детерминант позволит совершенствовать долгосрочные комплексные прогнозы развития территорий.

**Цель работы** — совершенствование методики выделения демографических региональных структур на примере анализа рождаемости населения Украины и формирующих ее факторов, а также статистическая оценка этого влияния.

Рабочая гипотеза исследования состояла в предположении, что социально-экономическое развитие регионов оказывает влияние на дифференциацию детородной активности населения.

Для достижения поставленной цели потребовалось решение следующих задач:

1) выявление оптимального и устойчивого разбиения регионов на группы;

2) оценка статистических взаимосвязей целевых и факторных признаков.

Для решения подобных задач (разбиение объектов на группы по совокупности признаков, выделение латентных факторов, статистическая оценка взаимосвязей) применялся набор многомерных статистических методов (кластерный, факторный, регрессионный анализ) [24, 25].

**Информационная база.** Информационная база исследования основывалась на показателях регионального развития Украины, предоставленных Институтом социальных и демографических исследований, рассчитанных по официальным статистическим данным [26]. Массив состоял из ежегодных статистических показателей 27 объектов статистической отчетности (24 области, АР Крым, города Киев и Севастополь) за 1999–2006 гг.

Целевые показатели, характеризующие детородную активность:

1) средний возраст матери при рождении ребенка;

2) средний возраст матери при рождении ребенка при стационарном населении;

3) суммарный коэффициент рождаемости — количество рождений на одну женщину (фертильность);

4) специальный коэффициент рождаемости — число рождений в среднем на 1000 женщин детородного возраста;

5) нетто-коэффициент воспроизводства женского поколения.

Состояние рынка труда характеризуется следующими показателями:

1) уровень экономической активности, женщины, %;

2) уровень занятости, женщины, %;

3) уровень безработицы, городское население, %;

4) удельный вес работающих в условиях неполного рабочего дня (недели) в общей численности занятых, %.

Материальное положение населения было представлено такими показателями:

1) среднемесячная заработная плата;

2) средний уровень совокупных затрат в расчете на одного человека;

3) соотношение совокупных расходов и совокупных доходов;

- 4) разница между денежными расходами и денежными доходами;
- 5) доля денежных доходов в общих доходах;
- 6) уровень бедности;
- 7) глубина бедности;
- 8) коэффициент концентрации дохода Джини;
- 9) количество минимальных потребительских корзин, которые можно приобрести на сумму среднемесячного дохода (по региональным ценам);
- 10) обеспеченность личными автомобилями (на 100 семей);
- 11) доля расходов на питание в совокупных расходах;
- 12) доля работающих, которые получают заработную плату ниже минимальной, %.

Показатели, характеризующие социальную среду:

- 1) доля лиц с высшим образованием в возрасте 25–60 лет, %;
- 2) коэффициент преступности;
- 3) соотношение разводов и браков.

Показатели макроэкономической ситуации:

- 1) ввод жилых домов (тыс. кв. м. общей площади),
- 2) удельный вес населения, которое проживает в городах с численностью свыше 100 тыс.,
- 3) объем расходов местных бюджетов на социальную сферу (грн., в расчете на одного человека).

Таким образом, состояние социально-демографических объектов (областей Украины) задано набором матриц "объект — признак":

$$X(t_l) = \begin{bmatrix} x_{11}(t_l) & x_{12}(t_l) & \dots & x_{1n}(t_l) \\ x_{21}(t_l) & x_{22}(t_l) & \dots & x_{2n}(t_l) \\ \dots & \dots & \dots & \dots \\ x_{m1}(t_l) & x_{m2}(t_l) & \dots & x_{mn}(t_l) \end{bmatrix}, \quad (1)$$

где  $x_{ij}(t)$  —  $j$ -й признак  $i$ -го объекта;  $i = 1, \dots, m$ ,  $m$  — количество объектов ( $m = 27$ );  $j = 1, \dots, n$ ,  $n$  — количество целевых и факторных признаков,  $n = 27$ ;  $t_l$  — время (годы),  $l = 1, \dots, L$ ;  $L = 8$  (1999–2006 гг.)

#### Методы исследования

На первом этапе исследования решалась задача разбиения объектов на непересекающиеся подмножества  $C_k$ , (кластеры), количество которых заранее неизвестно  $k = 1, \dots, K$ . Выделение кластеров проводилось по набору целевых признаков — пять показателей рождаемости.

Решение проблемы поиска оптимальной и устойчивой структуры в динамических данных потребовало изменений в традиционной схеме кластерного анализа.

В последнее время развивается индексный подход к оценке качества кластерного решения. Разработаны инструменты оценки структуры кластеров, которые применяются для вычисления оптимального количества разбиений при кластеризации различными алгоритмами [27–29]. Индексный подход применялся нами для поиска устойчивой кластеризации областей Украины по показателям депопуляции [30].

Разработанная нами схема выявления оптимальной и устойчивой кластерной структуры состоит из следующих этапов:

1. Стандартизация показателей.

Демографические и социально-экономические показатели отличаются размерностью, поэтому исходные данные нормируют

$$Zscore(x_{ij}) = \frac{(x_{ij} - \bar{x}_j)}{S_j}, \quad (2)$$

где  $\bar{x}_j$  — вектор средних значений признаков,  $S_j$  — стандартные отклонения.

2. Кластеризация во временном разрезе:

– кластеризация иерархическими методами (для каждого года определяется оптимальное количество кластеров).

– кластеризация плоскими методами (проводятся многократные разбиения с использованием нескольких методов и ряда критериев для определения оптимального количества кластеров).

3. Сравнение результатов кластеризации по каждому году (рассчитываются внешние индексы разбиений, полученных с применением различных алгоритмов).

4. Кластеризация массива данных за все годы по сходству динамики признаков.

5. Проверка на устойчивость кластерной структуры (сравнение состава кластеров, полученных разными методами кластеризации и способами структуризации исходных данных).

Расчеты проводились с применением пакетов STATISTICA, MATLAB.

Применялись методы иерархической кластеризации (одиночной связи, центроидный метод, средней связи, полной связи, медианный метод, метод Уорда). Мера близости — евклидово расстояние, квадрат евклидова расстояния.

Плоские разбиения проводились алгоритмами:  $k$ -средних ( $k$ -means), медианным (PAM), а также методом самоорганизующихся карт Кохонена (SOM).

Внутренние индексы (критерии) характеризуют степень разброса внутри кластеров и между ними (компактность и отделимость кластеров). По определению более качественным считается разбиение на такое количество кластеров, при котором достигается оптимальное значение индексной функции (при условии хорошей интерпретации результатов). Рассчитывались: индексы RMSSTD, SPR, RS — для иерархических методов; для всех методов — индекс Дана (D), “ширина силуэта”(Kauffmann), индекс Дэвиса-Болдуина (DB), мера Калински-Гарабача (CH), индекс Krzanowski-Lai (KL) [27, 28].

Внешние индексы показывают степень совпадения состава кластеров в разбиениях, полученных разными способами. Рассчитывались индексы Рэнда (Rand), Adjusted Rand, Жаккарда (Jaccard), Фолка-Меллоу (Fowlkes-Mallows) [27, 28].

## Основные результаты кластеризации показателей рождаемости

В разбиениях, полученных иерархическими методами, оптимальное количество кластеров, при котором достигались максимумы индексов СН и KL, для каждого года было равно четырем. Не для каждого временного среза можно было точно определить оптимальное количество кластеров по функции RMSSTD. На оптимальное количество кластеров, равное четырем, всегда указывал максимум индекса SPR.

Алгоритм полной связи можно выделить как наиболее разделяющий, так как его разбиения имели большие значения индекса «ширина силуэта», чем полученные другими методами кластеры. Одним из лучших иерархических методов был метод Уорда, который выделял структуры, подобные полученным плоскими методами ( $k$ -средних).

Для всех временных срезов при кластеризации разделяющими методами оптимум внутренних индексов соответствовал четырем кластерам. Методом SOM выделены начиная с 2002 г. из кластера «Западный регион» в отдельный кластер Ивано-Франковская и Закарпатская и области.

Для каждого года проводилось сравнение разбиений, полученных разными методами (по величине внешних индексов). Максимальное значение внешних индексов достигалось при количестве кластеров равном  $K = 4$ . Ошибка несовпадения состава кластеров изменялась в диапазоне 2–10 %.

График объединения в кластеры (дендрограмма) областей Украины по сходству динамики пяти показателей рождаемости, рассчитанный по методу Уорда, представлен на рис. 1 с метками номеров основных кластеров. В кластер 1 вошли города Киев и Севастополь, Одесская область, АР Крым. Кластер 2 формируют области Центра и Востока Украины. В состав кластера 3 вошли области Западного региона. Кластер 4 сформирован из Волынской и Ровенской областей.

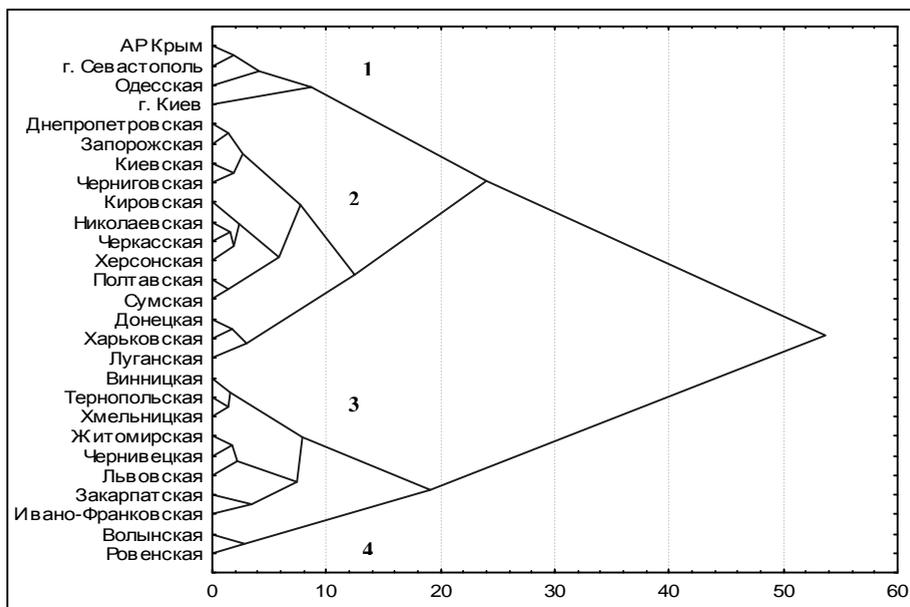


Рис. 1 Дендрограмма кластеризации показателей рождаемости (1999–2006 гг.), метод Уорда

Дисперсионный анализ подтвердил совпадение средних значений в четырех группах, рассчитанных методами  $k$ -средних и Уорда.

Вычислялись внутренние индексы этого разбиения, полученного кластеризацией по сходству динамики показателей рождаемости. Расчеты индексов (ширина силуэта, Данна, СН, KL) показали, что наиболее компактными и обособленными являются два кластера, один из которых объединяет кластер 1 и кластер 2, а второй — 3-й и 4-й.

Таким образом, кластерная структура, состоящая из четырех групп регионов, является устойчивой, так как она выделяется различными алгоритмами и различными способами сравнения исходных данных (по каждому году, по динамике, по всему массиву), а также удовлетворяет критериям качества кластеризации.

На рис. 2 представлена динамика средних значений показателей рождаемости в каждом кластере: коэффициента суммарной рождаемости (рис. 2, а) и возраста матери (рис. 2, б).

Как видим на рис. 2, а, б, для всех кластеров характерна общая тенденция восстановления сниженной рождаемости и увеличение возраста матери.

Характерные признаки четырех кластеров:

Кластер 1 — сочетание высокого возраста матери при рождении ребенка и сравнительно небольшой рождаемости.

Кластер 2 — самая низкая рождаемость (в него вошло большинство областей центрального и восточного региона), нетто-коэффициент воспроизводства и возраст матери ниже среднего.

Кластер 3 — рождаемость, близкая к среднему по Украине (западные области).

Кластер 4 — самая высокая рождаемость (Волинская, Ровенская области).

Подробнее типология рождаемости с учетом динамики показателей рождаемости нами описаны в работе [31].

### **Факторный анализ показателей регионального развития**

Следующим шагом стал поиск факторов, которые привели к изменению региональной рождаемости. Для этого применяли методы факторного анализа [24, 25].

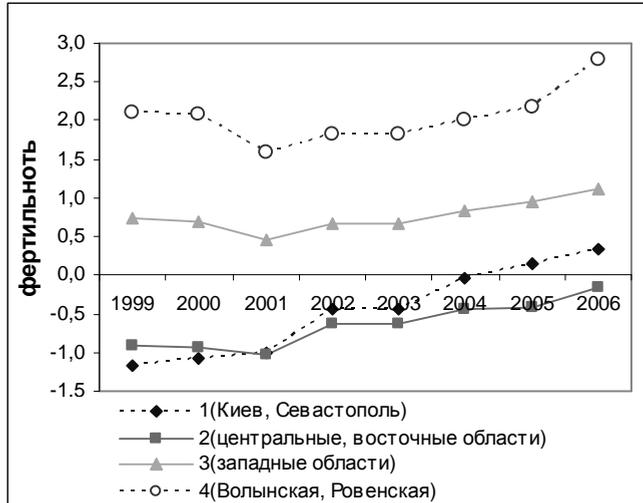
При помощи метода главных компонент сократили количество социально-экономических показателей, характеризующих региональное развитие в 1999–2006 гг. до небольшого числа общих или латентных факторов.

Первичные показатели можно представить в виде линейной комбинации главных факторов:

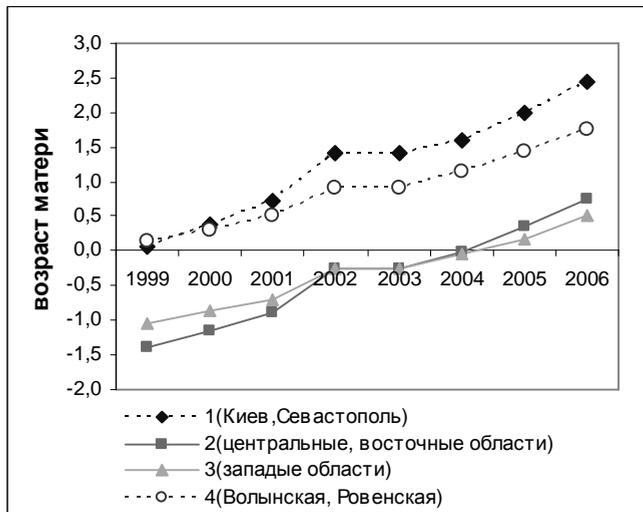
$$x_i = \sum_{j=1}^k a_{ij} F_j + l_i, \quad (3)$$

где  $x_i$  — факторные признаки ( $i = 1, p$ ),  $p = 22$ ,  $a_{ij}$  — факторные нагрузки,  $F_j$  — латентные факторы,  $k$  — количество факторов,  $l_i$  — остатки.

Процедура факторного анализа осуществлялась по корреляционной матрице совокупности (22) первичных факторных признаков. Предварительно определялась возможность применения факторного анализа по критерию сферичности Барлетта, а также критерию адекватности выборки Кайзера-Мейера-Олкина.



*a*



*b*

Рис. 2. Динамика стандартизованных показателей рождаемости в кластерах: *a* — суммарная рождаемость, *b* — возраст матери

Были выделены и отобраны по критериям Кайзера и Кэттеля четыре главные компоненты. В табл. 1 представлены собственные значения и дисперсии, которые объясняют выделенные компоненты (факторы). Необходимо отметить, что включение пятого и шестого факторов незначительно увеличивает процент объясненной дисперсии (на 6,3 % и 5,6 % соответственно), однако затрудняет последующую интерпретацию результатов. Поэтому в дальнейшем была использована 4-факторная модель.

Выделенные четыре фактора объясняют 61,24 % дисперсии первичных факторных признаков.

Факторная структура корреляционной матрицы уточнялась по методу ортогонального вращения варимакс (Varimax). В табл. 2 приведена факторная матрица после вращения, в которую включены переменные с высокими абсолютными значениями факторных нагрузок.

**Таблица 1**

*Выделенные компоненты (факторы)*

Компонента	Собственное значение	Дисперсия, %	Накопленная дисперсия, %
1	6,29	28,60	28,60
2	3,33	15,15	43,75
3	2,26	10,28	54,03
4	1,59	7,21	61,24

При анализе матрицы факторных нагрузок (табл. 2) была выделена следующая факторная структура.

Первая компонента, объединяющая восемь показателей, имеет высокие факторные нагрузки на показатель среднемесячной заработной платы и некоторые показатели материального положения (отрицательные значения нагрузок на показатели, характеризующие занятость населения). Это позволяет интерпретировать первый фактор, который объясняет 28,6 % общей дисперсии как «материальное положение».

Второй фактор объединяет социальные показатели: удельный вес населения, проживающего в городах численностью более 100 тыс., коэффициент преступности, долю лиц с высшим образованием среди населения старше 25 лет, соотношение разводов и браков. Этот фактор, объясняющий 15,15 % общей дисперсии, можно определить как «социально-поселенческий».

Третий фактор характеризует уровень обеспеченности в регионах, объединяя несколько показателей экономического неравенства (фактор «неравенство»).

Четвертый фактор характеризует занятость женщин («женская занятость»).

Таким образом, влияние комплекса показателей регионального развития на рождаемость можно свести к действию факторов: 1) материальное положение, 2) социально-поселенческий, 2) бедность (экономическое неравенство), 4) женская занятость.

**Анализ кластерной структуры региональной рождаемости с учетом основных факторов.**

После формирования комплексных факторов, которые определяют влияние социально-экономического развития на репродуктивную активность населения регионов, была проведена повторная кластеризация с учетом двух основных факторов (материальное положение и социально-поселенческий фактор).

Сравнивали результаты двух кластеризаций: первая – разбиение по показателям рождаемости, состав которой был описан выше (см. рис. 1), и вторая — классификация по показателям рождаемости и двум основным факторам. Совпадение состава четырех кластеров этих разбиений было почти

полным (Rand = 0,928; Adjusted Rand = 0,838; Jaccard = 0,805; Fowlkes-Mallows = 0,892), классификации отличались только принадлежностью Одесской области к другому кластеру. Эти результаты также подтверждают оптимальность и устойчивость выделенной кластерной структуры.

**Таблица 2**

*Матрица факторных нагрузок (после вращения)*

Показатель	Фактор			
	1	2	3	4
Средняя зарплата	0,94			
Совокупные расходы	-0,84			
Соотношение совокупных расходов и доходов	0,88			
Разница между расходами и доходами	-0,78			
Доля денежных доходов в общих доходах	0,79			
Доля расходов на питание	-0,80			
Уровень безработицы	-0,76			
Удельный вес неполной занятости	-0,74			
Доля городского населения		0,81		
Преступность		0,85		
Уровень образования		0,41		
Соотношение разводов и браков		0,74		
Уровень бедности			-0,81	
Глубина бедности			-0,59	
Обеспеченность автомобилями			0,59	
Ввод жилых домов			0,50	
Экономическая активность женщин				0,97
Женская занятость				0,94

Визуализация результатов (проекции центров кластеров на плоскость «показатель рождаемости — фактор») помогает оценить постоянство выявленной кластерной структуры в данных региональной рождаемости с 1999 по 2006 г.

На рис. 3, а–г представлены траектории центров кластеров на плоскость «показатель рождаемости (фертильность, возраст матери) — материальное положение».

Изучаемый период соответствует выходу из социально-экономического кризиса 90-х гг. Во всех кластерах за фазой падения рождаемости следует фаза восстановления. Вслед за ростом фактора 1 (улучшением материального положения, которое началось в 2000 г.) произошло некоторое возрастание фертильности и возраста матери при рождении ребенка.

Наиболее длительной фаза снижения рождаемости была в кластерах 4 (Волынская, Ровенская области) и 3 (западные области Украины).

Необходимо отметить, что кластеры изначально отличаются исходными уровнями фактора 1 (материальное положение), и со временем это различие сохраняется. Можно предположить, что именно это определяет различие длительностей и глубины фазы снижения фертильности — отложенных рождений. Наибольшим по сравнению с другими кластерами был прирост рождаемости в кластере 1 (С1 — Киев, АР Крым, Севастополь, Одесская область), для которого значения фактора 1 (материальное положение) были исходно большими.

Ближние по исходным уровням рождаемости кластеры С1 и С2 существенно отличаются по возрасту матери. По-видимому, это

свидетельствует о переходе в наиболее урбанизированном кластере С1 на общеевропейскую модель отложенных первых рождений, что проявляется в увеличении возраста матери [14]. Траектории, представляющие зависимость специального коэффициента рождаемости от фактора 1 (рис 3, в), в основном повторяют изменения траекторий суммарного коэффициента рождаемости в каждом кластере.

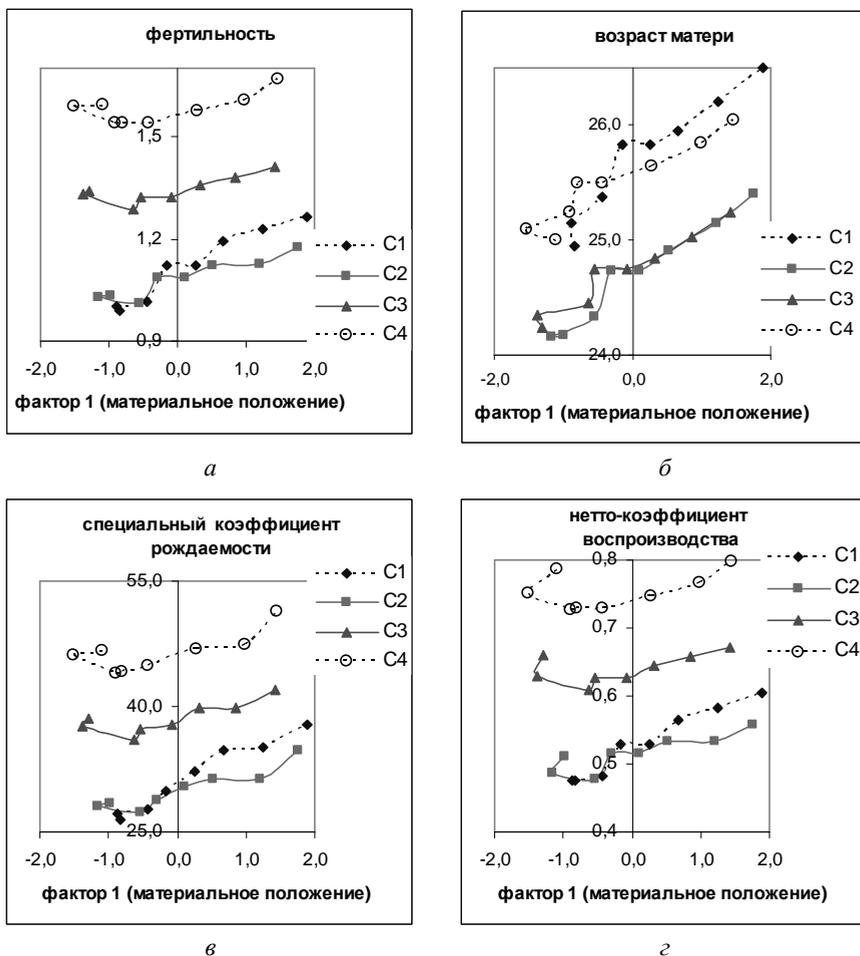


Рис.3. Проекция траекторий кластерных центров на плоскость кластеров (1999–2006 гг.)

На рис 3, г представлены графики траекторий кластерных центров, а именно их проекции на плоскость нетто-коэффициента воспроизводства женского населения и фактора 1. Нетто-коэффициент воспроизводства количественно оценивает меру замещения материнского поколения дочерним. Как видим на рис. 3, г, нетто-коэффициент в 1999–2006 гг. изменялся в диапазоне 0,45–0,8, что означает далеко не полное замещение поколения матерей (от 45 % до 80 %). Даже в кластере 4 с самым высоким уровнем детородной активности в Украине (С4 — Волынская и Ровенская области) произошло обвальное падение воспроизводства женского населения, как за счет снижения рождаемости, так и за счет роста смертности женского населения. В этом кластере только в 2006 г. уровень замещения

женского поколения достиг 80 % и стал близок к исходному уровню 1999 г.

Как известно, социально-поселенческий фактор (тип поселения, степень урбанизации) во многом детерминирует уровень рождаемости [22, 23, 30]. Необходимо отметить, что выделенные кластеры отличаются между собой по уровням фактора 2 (урбанизация).

На рис. 4, *а* приведены траектории кластерных центров в плоскости фертильность — фактор 2, а на рис. 4, *б* фертильность — фактор 3. Как можно увидеть на рис. 4, *а*, диапазон изменений фактора 2 (урбанизация) для каждого из кластеров был незначительным. Кластер 1 отличается от других высоким уровнем урбанизации (фактор 2) и неравенства (фактор 3). Отметим, что некоторое уменьшение значений фактора 2, которое наблюдалось в конце периода 1999–2006 гг. в кластерах 1 и 2, имеющих высокий исходный уровень урбанизации, произошло из-за уменьшения соотношения разводов и браков, а также уровня преступности.

Дифференциация кластеров сохранялась также и по фактору 3 (неравенство в обеспеченности). Кластер 1 отличается от других наибольшим уровнем этого фактора (что соответствует относительно меньшей бедности, большей обеспеченности автомобилями, жильем) (рис. 4).

Степень дифференциации четырех типологических кластеров с годами меняется, что подтверждается результатами дисперсионного анализа по годам и кластерам. В рассматриваемом периоде по показателям детородной активности регионов и фактору «материальное положение» сохраняется статистически достоверное различие между средними значениями этих показателей в четырех кластерах.

Таким образом, в массиве данных регионального развития Украины за 1999–2006 гг. выделена устойчивая и оптимальная кластерная структура, состоящая из четырех типологических кластеров, которые отличаются как показателями репродуктивной активности населения, так и уровнями комплексных социально-экономических факторов.

Разработанная методика выделения устойчивых демографических региональных структур использует гибридный подход (кластерные и факторные методы анализа), индексный метод обоснования результата кластеризации, кроме того, дополнена несколькими способами структуризации исходных данных. Вычислительная схема включает методы кластерного и факторного анализа, расчеты индексных функций качества кластеризации, несколько способов структуризации исходных данных.

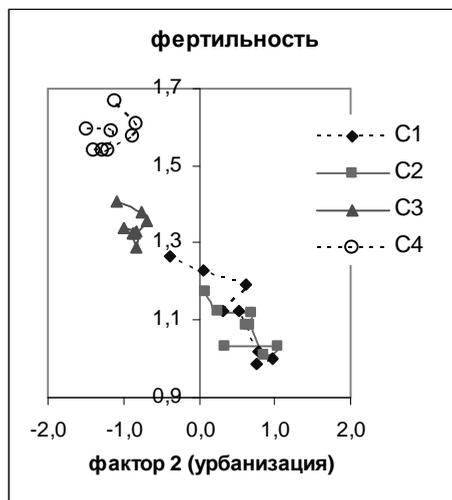
### **Регрессионная модель зависимости рождаемости от основных факторов**

Для получения статистической оценки относительного влияния основных факторов на рождаемость был проведен множественный регрессионный анализ.

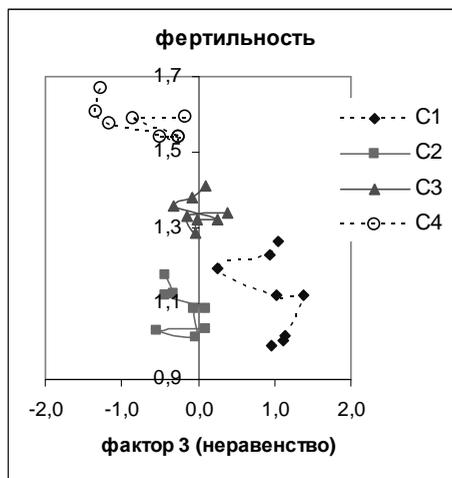
В линейной форме уравнение зависимости рождаемости от выделенных четырех основных факторов имеет вид:

$$Fert_i(t) = b_0 + b_1 \cdot Fac1_i(t) + b_2 \cdot Fac2_i(t) + b_3 \cdot Fac3_i(t) + b_4 \cdot Fac4_i(t) + \varepsilon(t), \quad (4)$$

где  $i = 1, \dots, m$  — соответствует количеству объектов социально-демографической отчетности,  $m = 27$ ;  $t$  — периоду времени (с 1999 по 2006 г.);  $Fert$  — коэффициент суммарной рождаемости (фертильность) на уровне региона;  $b_1, b_2, b_3, b_4$  — коэффициенты регрессии;  $Fac1$  — материальное положение (благополучие),  $Fac2$  — социально-поселенческий (урбанизация),  $Fac3$  — неравенство в обеспеченности,  $Fac4$  — женская занятость,  $\varepsilon$  — неучтенные факторы.



а



б

Рис. 4. Траектории кластерных центров

В табл. 3 приведены результаты расчетов 4-факторной модели фертильности, рассчитанной по данным всех регионов Украины (1999–2006 гг). Значение коэффициента множественной корреляции  $R = 0,809$  ( $F = 98,2$ ,  $p < 0,00001$ ). Зависимость  $Fert$  от  $Fac1, Fac2, Fac3, Fac4$  можно полагать достаточно тесной, 65,5 % вариации фертильности определяются изменениями четырех независимых факторов. Остальная часть объясняется влиянием неучтенных нами факторов. Коэффициенты модели (4) рассчитывались по методу наименьших квадратов. Значения коэффициентов регрессии  $b(i)$

статистически значимы, уровень значимости для  $b_4$  несколько ниже.

В соответствии со значениями коэффициентов модели (4) фертильность прямо пропорционально зависит от изменения фактора 1 (материального благополучия —  $FAC1$ ). Фактор 2 (социально-поселенческий, урбанизация,  $FAC2$ ) влияет с противоположным знаком, фактор 3 (неравенство в обеспеченности,  $FAC3$ ) также имеет значимое отрицательное влияние на показатель детородной активности населения.

**Таблица 3**

*Параметры регрессионной модели зависимости рождаемости от комплексных социально-экономических факторов*

Фактор	$\beta$ -коэффициент	Статистическая ошибка $\beta$	Коэффициент регрессии $b(i)$	Статистическая ошибка $b(i)$	Коэффициент Стьюдента $t(266)$	Уровень значимости $p$
Постоянная	—	—	1,202	0,0075	159,6	0,0000
$FAC1$	0,119	0,0407	0,022	0,0075	2,9	0,0030
$FAC2$	-0,763	0,0407	-0,141	0,0075	-18,7	0,0000
$FAC3$	-0,222	0,0407	-0,041	0,0075	-5,4	0,0000
$FAC4$	-0,066	0,0407	-0,012	0,0075	-1,6	0,1053

Анализ частных  $\beta$ -коэффициентов приводит к выводу, что наибольшее влияние на дифференциацию детородной активности имеет фактор 2 (социально-поселенческий,  $FAC2$ ), так как для этого фактора абсолютное значение  $\beta$ -коэффициента больше, чем для других факторов ( $\beta_2 = -0,76$ ). Отметим, что степень влияния фактора 3 (неравенство,  $\beta_3 = -0,22$ ) сопоставима с влиянием фактора «материальное положение» ( $\beta_1 = -0,119$ ).

Необходимо отметить, что ведущий фактор 1 (по величине объясненной им дисперсии) претерпевал наибольшие изменения с годами по сравнению с другими, при этом степень различия между кластерами по величине этого фактора с годами изменялась незначительно. Как уже отмечалось, уровни рождаемости в большей степени зависят от фактора 2 (социально-поселенческий). Поэтому с годами сохранялось соотношение уровней региональной рождаемости (между кластерами).

С целью уточнения взаимосвязи фертильности и основных факторов, нами проведено исследование модели (4) для каждого объекта из четырех типологических кластеров. Остатки между предсказанными по модели (4) и наблюдаемыми значениями (погрешность модели) не превышала 10 %. Больше всего эта погрешность оказалась для Ровенской области (6,2 %), а именно рассчитанные величины были несколько меньше наблюдаемых значений фертильности.

Для областей, входящих в кластер 1 (урбанистический), влияние изменений факторов «материальное положение» ( $\beta_1 = 0,597$ ) и «женская занятость» ( $\beta_4 = -0,47$ ) больше, чем двух других факторов. Для иллюстрации этой зависимости, близкой к линейной, приведен график на рис. 5. Здесь приведены наблюдаемые значения показателя фертильности для г. Киева (1999–2006 гг.) и значения рассчитанные по 4-факторной модели.

Для кластера 3 (Волинская и Ровенская области) характерна наиболее тесная взаимосвязь между фертильностью и фактором 4 (экономическая активность женщин –  $\beta_4 = -1,32$ ). На рис. 6 приведены графики взаимосвязи фертильности и экономической активности женщин для Волинской области, наблюдаемые значения и значения, рассчитанные по модели (5).

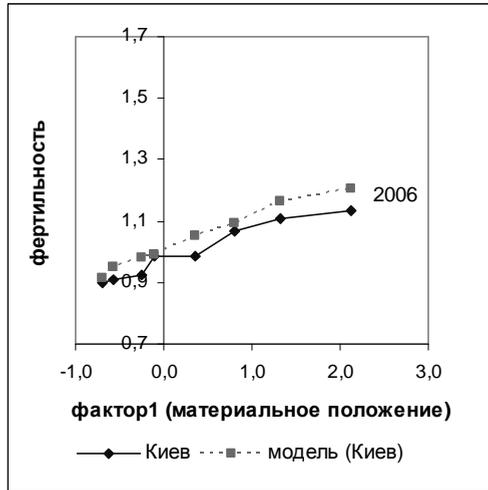


Рис. 5. Зависимость фертильности (г. Киев) от изменения материального благополучия, фактор 1 (1999–2006 гг.)

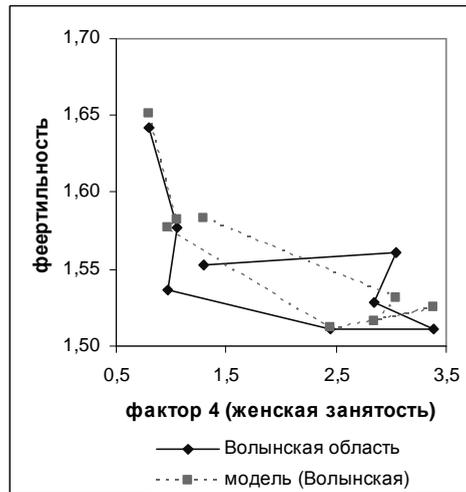


Рис. 6. Изменение фертильности (Волинская область) в зависимости от экономической активности женщин (фактор 4)

## Выводы

1. Разработанная модифицированная схема кластерного анализа, которая включает несколько методов и способов формирования признаков кластеризации, расчет индексов качества кластеризации, позволила выявить устойчивую кластерную структуру региональной рождаемости населения Украины в 1999–2006 гг.

2. Проведенный факторный анализ показателей регионального развития областей позволил выделить четыре комплексных фактора: материальное положение населения, особенности его размещения, бедность, экономическая активность женщин, от изменения этих факторов зависит дифференциация репродуктивной активности населения Украины.

3. Разработанная множественная линейная регрессионная модель зависимости фертильности от основных факторов позволила дать сравнительную оценку их влияния на уровень репродуктивной активности населения. Модель дает возможность объяснить устойчивость кластерной структуры региональной рождаемости детерминирующим влиянием социально-поселенческого фактора (урбанизация), изменения которого были незначительными по сравнению с материальным фактором.

1. *Becker G.* A treatise on the family. — Cambridge: Harvard University press, 1991. — 288 p.
2. *Schultz T.* An economic model of family planning and fertility // *Journal of Political Economy.* — 1969. — № 77 (2). — P. 153–180.
3. *Easterlin R.* An Economic Framework for Fertility Analysis // *Studies in Family Planning.* — 1975. — Vol. 6. — N 3. — P. 54–63.
4. *Easterlin R.A., Pollak R.A., Wachter M.L.* Towards a More General Economic Model of Fertility Determination: Endogenous Preferences and Natural Fertility // *Population and Economic Change in Developing Countries.* — Chicago: University of Chicago Press, 1980. — P. 81–150.
5. *Becker G., Barro R.* Altruism and the Economic Theory of Fertility // *Population and Development Review.* — 1986. — Vol. 12, Supplement: Below-Replacement Fertility in Industrial Societies: Causes, Consequences, Policies. — P. 69–76.
6. *Warren C.R.* The Economic Theory of Fertility Over Three Decades // *Population Studies.* — 1997. — Vol. 51, N 1. — P. 63–74.
7. *Myrskylä M., Kohler H.P., Billari F.* Advances in development reverse fertility declines // *Nature.* — 2009. — 460 (6). — P. 741–743.
8. Демографічна криза в Україні. Проблеми дослідження, витоки, складові, напрями протидії / за ред. В. Стешенко.— Київ: Інститут економіки НАН України, 2001. — 422 с.
9. *Курило І.О.* Про чинники сучасного зниження народжуваності в Україні / Демографічні дослідження. — К.: Ін-т економіки НАН України, 2001. — Вип. 23. — С. 7–21.
10. *Мандибуря В.О.* Падіння рівня життя як чинник демографічної кризи в Україні // Наукові записки КНУ ім. Тараса Шевченка. — К.: КПВД Педагогіка, 2004. — 4. — С. 92–103.
11. *Лин Д.Г., Предыбайло С.Д.* Демографическая ситуация в современной Беларуси // *Народонаселение.* — 2005. — № 1. — С. 104–118.
12. *Палій О.М.* Демографічна ситуація в країнах Центральної та Східної Європи // *Демографія та соціальна економіка.* — 2007. — № 2. — С. 3–13.
13. *Ревун В.И.* Демографическое развитие Беларуси, России и Украины в условиях депопуляции: Автореф. дис. ... д-ра эк. наук: 08.00.05. — М., 2009. — 46 с.
14. *Курило І.О.* Тенденції народжуваності в Україні у контексті подальшої трансформації суспільства. Демографія та соціальна економіка. — 2006. — № 1. — С. 38–45.
15. *Населення України. Народжуваність в Україні у контексті суспільно-трансформаційних процесів.* — К.: АДЕФ-Україна, 2008. — 288 с.
16. *Скрябина Я.А.* Особенности репродуктивного поведения населения современной России // *Инновации и инвестиции.* — 2011. — №1. — С. 86–90.
17. *Халкечев М.Н.* Демографическая дифференциация регионов России: динамика и развитие. — М.: ЦСП, 2006. — 224 с.

18. Шевчук П. Україна: мультирегіональна модель народжуваності // Регіональна економіка. — 2000. — № 2. — С.128–133.
19. Населення України — 2004. Регіональні аспекти демографічного розвитку / С.І. Пирожков, О.Г. Рогожин, П.С. Шевчук та ін. — К.: ІДСД НАНУ, Держкомстат України, 2004. — 342 с.
20. Людський розвиток регіонів України: аналіз та прогноз (колективна монографія) / за ред. Е.М. Лібанової. — К.: Ін-т демографії та соціальних досліджень НАН України, 2007. — 328 с.
21. Людський розвиток в Україні: мінімізація соціальних ризиків (колективна науково-аналітична монографія) / За ред. Е.М. Лібанової. — К.: Ін-т демографії та соціальних досліджень ім. М.В.Птухи НАН України, Держкомстат України, 2010. — 496 с.
22. Ильишев А.М., Багирова А.П. Факторы репродуктивной активности населения: анализ региональной дифференциации // Региональная экономика и социология. — 2009. — № 1. — С. 92–110.
23. Бутс Б., Дробышевский С., Кочеткова О. Типология российских регионов — М.: Институт экономики переходного периода, СЕРРА, 2002. — 159 с.
24. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы. — М.: Финансы и статистика, 1998. — 284 с.
25. Факторный, дискриминантный и кластерный анализ. — М.: Финансы и статистика, 1989. — 215 с.
26. Державний комітет статистики України. — [www.ukrstat.gov.ua](http://www.ukrstat.gov.ua).
27. Milligan G., Cooper, M. An examination of procedures for determining the number of clusters // Psychometrika. — 1985. — 50. — P. 159–179.
28. Halkidi M., Batistakis V. On Clustering Validation Techniques // Journal of Intelligent Information Systems. — 2001. — 17, N 2, 3. — P. 107–145.
29. Wang K., Wang B., Peng L. CVAP: Validation for cluster analyses // Data Science Journal. — 2009. — № 8. — P. 88–93.
30. Кривова О.А., Курило І.А. Регіональна типологія естественного прироста населення України: кластерний підхід // Кибернетика и вычислительная техника. — 2011. — Вып. 164. — С. 89–102.
31. Кривова О.А., Коваленко О.С., Курило І.А. Регіональні особливості народжуваності в Україні та їх детермінанти (досвід аналізу багатомірними статистичними методами) // Регіональні аспекти розвитку продуктивних сил України: Зб. наук. праць. — Тернопіль: Економічна думка, 2009. — Вип. 14. — С. 44–47.

Международный научно-учебный центр  
информационных технологий и систем  
НАН Украины и Министерства образования  
и науки, молодежи и спорта Украины, Киев  
Институт демографии и социальных  
исследований имени В.М. Птухи НАН Украины

Получено 08.08.2012